

1639



RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/800,770

DATE: 01/14/2003

TIME: 14:55:03

Input Set : A:\A68614-1.ST25.txt

Output Set: N:\CRF4\01142003\I800770.raw

3 <110> APPLICANT: Kinsella, Todd
 5 <120> TITLE OF INVENTION: IN VIVO PRODUCTION OF CYCLIC PEPTIDES
 7 <130> FILE REFERENCE: A-68614-1/DJB/RMS/RMK
 9 <140> CURRENT APPLICATION NUMBER: US 09/800,770
 10 <141> CURRENT FILING DATE: 2001-03-06
 12 <150> PRIOR APPLICATION NUMBER: US 60/187,130
 13 <151> PRIOR FILING DATE: 2000-03-06
 15 <160> NUMBER OF SEQ ID NOS: 90
 17 <170> SOFTWARE: PatentIn version 3.1
 19 <210> SEQ ID NO: 1
 20 <211> LENGTH: 431
 21 <212> TYPE: PRT
 22 <213> ORGANISM: Synechocystis PCC6803
 24 <400> SEQUENCE: 1
 26 Gly Cys Ile Ser Gly Asp Ser Leu Ile Ser Leu Ala Ser Thr Gly Lys
 27 1 5 10 15
 30 Arg Val Ser Ile Lys Asp Leu Leu Asp Glu Lys Asp Phe Glu Ile Trp
 31 20 25 30
 34 Ala Ile Asn Glu Gln Thr Met Lys Leu Glu Ser Ala Lys Val Ser Arg
 35 35 40 45
 38 Val Phe Cys Thr Gly Lys Lys Leu Val Tyr Ile Leu Lys Thr Arg Leu
 39 50 55 60
 42 Gly Arg Thr Ile Lys Ala Thr Ala Asn His Arg Phe Leu Thr Ile Asp
 43 65 70 75 80
 46 Gly Trp Lys Arg Leu Asp Glu Leu Ser Leu Lys Glu His Ile Ala Leu
 47 85 90 95
 50 Pro Arg Lys Leu Glu Ser Ser Ser Leu Gln Leu Met Ser Asp Glu Glu
 51 100 105 110
 54 Leu Gly Leu Leu Gly His Leu Ile Gly Asp Gly Cys Thr Leu Pro Arg
 55 115 120 125
 58 His Ala Ile Gln Tyr Thr Ser Asn Lys Ile Glu Leu Ala Glu Lys Val
 59 130 135 140
 62 Val Glu Leu Ala Lys Ala Val Phe Gly Asp Gln Ile Asn Pro Arg Ile
 63 145 150 155 160
 66 Ser Gln Glu Arg Gln Trp Tyr Gln Val Tyr Ile Pro Ala Ser Tyr Arg
 67 165 170 175
 70 Leu Thr His Asn Lys Lys Asn Pro Ile Thr Lys Trp Leu Glu Asn Leu
 71 180 185 190
 74 Asp Val Phe Gly Leu Arg Ser Tyr Glu Lys Phe Val Pro Asn Gln Val
 75 195 200 205
 78 Phe Glu Gln Pro Gln Arg Ala Ile Ala Ile Phe Leu Arg His Leu Trp
 79 210 215 220
 82 Ser Thr Asp Gly Cys Val Lys Leu Ile Val Glu Lys Ser Ser Arg Pro

ENTERED

RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/800,770

DATE: 01/14/2003

TIME: 14:55:03

Input Set : A:\A68614-1.ST25.txt

Output Set: N:\CRF4\01142003\I800770.raw

```

83 225          230          235          240
86 Val Ala Tyr Tyr Ala Thr Ser Ser Glu Lys Leu Ala Lys Asp Val Gln
87          245          250          255
90 Ser Leu Leu Leu Lys Leu Gly Ile Asn Ala Arg Leu Ser Lys Ile Ser
91          260          265          270
94 Gln Asn Gly Lys Gly Arg Asp Asn Tyr His Val Thr Ile Thr Gly Gln
95          275          280          285
98 Ala Asp Leu Gln Ile Phe Val Asp Gln Ile Gly Ala Val Asp Lys Asp
99          290          295          300
102 Lys Gln Ala Ser Val Glu Glu Ile Lys Thr His Ile Ala Gln His Gln
103 305          310          315          320
106 Ala Asn Thr Asn Arg Asp Val Ile Pro Lys Gln Ile Trp Lys Thr Tyr
107          325          330          335
110 Val Leu Pro Gln Ile Gln Ile Lys Gly Ile Thr Thr Arg Asp Leu Gln
111          340          345          350
114 Met Arg Leu Gly Asn Ala Tyr Cys Gly Thr Ala Leu Tyr Lys His Asn
115          355          360          365
118 Leu Ser Arg Glu Arg Ala Ala Lys Ile Ala Thr Ile Thr Gln Ser Pro
119          370          375          380
122 Glu Ile Glu Lys Leu Ser Gln Ser Asp Ile Tyr Trp Asp Ser Ile Val
123 385          390          395          400
126 Ser Ile Thr Glu Thr Gly Val Glu Glu Val Phe Asp Leu Thr Val Pro
127          405          410          415
130 Gly Pro His Asn Phe Val Ala Asn Asp Ile Ile Val His Asn Ser
131          420          425          430
134 <210> SEQ ID NO: 2
135 <211> LENGTH: 200
136 <212> TYPE: PRT
137 <213> ORGANISM: Mycobacterium xenopi
139 <400> SEQUENCE: 2
141 Tyr Cys Ile Thr Gly Asp Ala Leu Val Ala Leu Pro Glu Gly Glu Ser
142 1          5          10          15
145 Val Arg Ile Ala Asp Ile Val Pro Gly Ala Arg Pro Asn Ser Asp Asn
146          20          25          30
149 Ala Ile Asp Leu Lys Val Leu Asp Arg His Gly Asn Pro Val Leu Ala
150          35          40          45
153 Asp Arg Leu Phe His Ser Gly Glu His Pro Val Tyr Thr Val Arg Thr
154          50          55          60
157 Val Glu Gly Leu Arg Val Thr Gly Thr Ala Asn His Pro Leu Leu Cys
158 65          70          75          80
161 Leu Val Asp Val Ala Gly Val Pro Thr Leu Leu Trp Lys Leu Ile Asp
162          85          90          95
165 Glu Ile Lys Pro Gly Asp Tyr Ala Val Ile Gln Arg Ser Ala Phe Ser
166          100          105          110
169 Val Asp Cys Ala Gly Phe Ala Arg Gly Lys Pro Glu Phe Ala Pro Thr
170          115          120          125
173 Thr Tyr Thr Val Gly Val Pro Gly Leu Val Arg Phe Leu Glu Ala His
174          130          135          140
177 His Arg Asp Pro Asp Ala Gln Ala Ile Ala Asp Glu Leu Thr Asp Gly

```

RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/800,770

DATE: 01/14/2003

TIME: 14:55:03

Input Set : A:\A68614-1.ST25.txt

Output Set: N:\CRF4\01142003\I800770.raw

```

178 145          150          155          160
181 Arg Phe Tyr Tyr Ala Lys Val Ala Ser Val Thr Asp Ala Gly Val Gln
182          165          170          175
185 Pro Val Tyr Ser Leu Arg Val Asp Thr Ala Asp His Ala Phe Ile Thr
186          180          185          190
189 Asn Gly Phe Val Ser His Asn Thr
190          195          200
193 <210> SEQ ID NO: 3
194 <211> LENGTH: 458
195 <212> TYPE: PRT
196 <213> ORGANISM: Chlamydomonas eugametos
198 <400> SEQUENCE: 3
200 Glu Cys Leu Thr Ser Asp His Thr Val Leu Thr Thr Arg Gly Trp Ile
201 1          5          10          15
204 Pro Ile Ala Asp Val Thr Leu Asp Asp Lys Val Ala Val Leu Asp Asn
205          20          25          30
208 Asn Thr Gly Glu Met Ser Tyr Gln Asn Pro Gln Lys Val His Lys Tyr
209          35          40          45
212 Asp Tyr Glu Gly Pro Met Tyr Glu Val Lys Thr Ala Gly Val Asp Leu
213          50          55          60
216 Phe Val Thr Pro Asn His Arg Met Tyr Val Asn Thr Thr Asn Asn Thr
217 65          70          75          80
220 Thr Asn Gln Asn Tyr Asn Leu Val Glu Ala Ser Ser Ile Phe Gly Lys
221          85          90          95
224 Lys Val Arg Tyr Lys Asn Asp Ala Ile Trp Asn Lys Thr Asp Tyr Gln
225          100          105          110
228 Phe Ile Leu Pro Glu Thr Ala Thr Leu Thr Gly His Thr Asn Lys Ile
229          115          120          125
232 Ser Ser Thr Pro Ala Ile Gln Pro Glu Met Asn Ala Trp Leu Thr Phe
233          130          135          140
236 Phe Gly Leu Trp Ile Ala Asn Gly His Thr Thr Lys Ile Ala Glu Lys
237 145          150          155          160
240 Thr Ala Glu Asn Asn Gln Gln Lys Gln Arg Tyr Lys Val Ile Leu Thr
241          165          170          175
244 Gln Val Lys Glu Asp Val Cys Asp Ile Ile Glu Gln Thr Leu Asn Lys
245          180          185          190
248 Leu Gly Phe Asn Phe Ile Arg Ser Gly Lys Asp Tyr Thr Ile Glu Asn
249          195          200          205
252 Lys Gln Leu Trp Ser Tyr Leu Asn Pro Phe Asp Asn Gly Ala Leu Asn
253          210          215          220
256 Lys Tyr Leu Pro Asp Trp Val Trp Glu Leu Ser Ser Gln Gln Cys Lys
257 225          230          235          240
260 Ile Leu Leu Asn Ser Leu Cys Leu Gly Asn Cys Leu Phe Thr Lys Asn
261          245          250          255
264 Asp Asp Thr Leu His Tyr Phe Ser Thr Ser Glu Arg Phe Ala Asn Asp
265          260          265          270
268 Val Ser Arg Leu Ala Leu His Ala Gly Thr Thr Ser Thr Ile Gln Leu
269          275          280          285
272 Glu Ala Ala Pro Ser Asn Leu Tyr Asp Thr Ile Ile Gly Leu Pro Val

```

RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/800,770

DATE: 01/14/2003

TIME: 14:55:03

Input Set : A:\A68614-1.ST25.txt

Output Set: N:\CRF4\01142003\I800770.raw

```

273      290      295      300
276 Glu Val Asn Thr Thr Leu Trp Arg Val Ile Ile Asn Gln Ser Ser Phe
277 305      310      315      320
280 Tyr Ser Tyr Ser Thr Asp Lys Ser Ser Ala Leu Asn Leu Ser Asn Asn
281      325      330      335
284 Val Ala Cys Tyr Val Asn Ala Gln Ser Ala Leu Thr Leu Glu Gln Asn
285      340      345      350
288 Ser Gln Lys Ile Asn Lys Asn Thr Leu Val Leu Thr Lys Asn Asn Val
289      355      360      365
292 Lys Ser Gln Thr Met His Ser Gln Arg Ala Glu Arg Val Asp Thr Ala
293      370      375      380
296 Leu Leu Thr Gln Lys Glu Leu Asp Asn Ser Leu Asn His Glu Ile Leu
297 385      390      395      400
300 Ile Asn Lys Asn Pro Gly Thr Ser Gln Leu Glu Cys Val Val Asn Pro
301      405      410      415
304 Glu Val Asn Asn Thr Ser Thr Asn Asp Arg Phe Val Tyr Tyr Lys Gly
305      420      425      430
308 Pro Val Tyr Cys Leu Thr Gly Pro Asn Asn Val Phe Tyr Val Gln Arg
309      435      440      445
312 Asn Gly Lys Ala Val Trp Thr Gly Asn Ser
313      450      455
316 <210> SEQ ID NO: 4
317 <211> LENGTH: 341
318 <212> TYPE: PRT
319 <213> ORGANISM: Chilo iridescent virus
321 <400> SEQUENCE: 4
323 Leu Cys Val Ala Pro Glu Thr Met Ile Leu Thr Glu Asp Gly Gln Phe
324 1      5      10      15
327 Pro Ile Lys Asp Leu Glu Gly Lys Ile Ile Lys Val Trp Asn Gly Asn
328      20      25      30
331 Glu Phe Ser Ser Val Thr Val Val Lys Thr Gly Thr Glu Lys Glu Leu
332      35      40      45
335 Leu Glu Val Glu Leu Ser Asn Gly Cys Thr Leu Ser Cys Thr Pro Glu
336      50      55      60
339 His Lys Phe Ile Ile Val Lys Ser Tyr Thr Glu Ala Lys Lys Gln Lys
340 65      70      75      80
343 Thr Asp Asp Asn Ala Ile Ala Asn Ala Glu Arg Val Asp Ala Gln Asp
344      85      90      95
347 Leu Lys Pro Arg Met Lys Leu Ile Lys Phe Asp Leu Pro Thr Leu Phe
348      100      105      110
351 Gly Asn Ser Glu His Asp Ile Lys Tyr Pro Tyr Thr His Gly Phe Phe
352      115      120      125
355 Cys Gly Asp Gly Thr Tyr Thr Lys Tyr Gly Lys Pro Gln Leu Ser Leu
356      130      135      140
359 Tyr Gly Asp Lys Lys Glu Leu Leu Thr Tyr Leu Asp Val Arg Thr Met
360 145      150      155      160
363 Thr Gly Leu Glu Asp Ala Ser Gly Arg Leu Asn Thr Trp Leu Pro Leu
364      165      170      175
367 Asp Leu Ala Pro Lys Phe Asp Val Pro Ile Asn Ser Ser Leu Glu Cys

```

RAW SEQUENCE LISTING

DATE: 01/14/2003

PATENT APPLICATION: US/09/800,770

TIME: 14:55:03

Input Set : A:\A68614-1.ST25.txt

Output Set: N:\CRF4\01142003\I800770.raw

```

368          180          185          190
371 Arg Met Glu Trp Leu Ala Gly Tyr Leu Asp Ala Asp Gly Cys Val Phe
372          195          200          205
375 Arg Asn Gly Thr Asn Glu Ser Ile Gln Val Ser Cys Ile His Leu Asp
376          210          215          220
379 Phe Leu Lys Arg Ile Gln Leu Leu Leu Ile Gly Met Gly Val Thr Ser
380 225          230          235          240
383 Lys Ile Thr Lys Leu His Asp Glu Lys Ile Thr Thr Met Pro Asp Gly
384          245          250          255
387 Lys Gly Gly Gln Lys Pro Tyr Ser Cys Lys Pro Ile Trp Arg Leu Phe
388          260          265          270
391 Ile Ser Ser Ser Gly Leu Tyr His Leu Ser Glu Gln Gly Phe Glu Thr
392          275          280          285
395 Arg Arg Leu Lys Trp Glu Pro Arg Gln Pro Gln Arg Asn Ala Glu Arg
396          290          295          300
399 Phe Val Glu Val Leu Lys Val Asn Lys Thr Gly Arg Val Asp Asp Thr
400 305          310          315          320
403 Tyr Cys Phe Thr Glu Pro Ile Asn His Ala Gly Val Phe Asn Gly Ile
404          325          330          335
407 Leu Thr Gly Gln Cys
408          340
411 <210> SEQ ID NO: 5
412 <211> LENGTH: 473
413 <212> TYPE: PRT
414 <213> ORGANISM: Candida tropicalis
416 <400> SEQUENCE: 5
418 Gly Cys Phe Thr Lys Gly Thr Gln Val Met Met Ala Asp Gly Ala Asp
419 1          5          10          15
422 Lys Ser Ile Glu Ser Ile Glu Val Gly Asp Lys Val Met Gly Lys Asp
423          20          25          30
426 Gly Met Pro Arg Glu Val Val Gly Leu Pro Arg Gly Tyr Asp Asp Met
427          35          40          45
430 Tyr Lys Val Arg Gln Leu Ser Ser Thr Arg Arg Asn Ala Lys Ser Glu
431          50          55          60
434 Gly Leu Met Asp Phe Thr Val Ser Ala Asp His Lys Leu Ile Leu Lys
435 65          70          75          80
438 Thr Lys Lys Asp Val Lys Ile Ala Thr Arg Lys Ile Gly Gly Asn Thr
439          85          90          95
442 Tyr Thr Gly Val Thr Phe Tyr Val Leu Glu Lys Thr Lys Thr Gly Ile
443          100          105          110
446 Glu Leu Val Lys Ala Lys Thr Lys Val Phe Gly His His Ile His Gly
447          115          120          125
450 Gln Asn Gly Ala Glu Glu Lys Ala Ala Thr Phe Ala Ala Gly Ile Asp
451          130          135          140
454 Ser Lys Glu Tyr Ile Asp Trp Ile Ile Glu Ala Arg Asp Tyr Val Gln
455 145          150          155          160
458 Val Asp Glu Ile Val Lys Thr Ser Thr Thr Gln Met Ile Asn Pro Val
459          165          170          175
462 His Phe Glu Ser Gly Lys Leu Gly Asn Trp Leu His Glu His Lys Gln

```

VERIFICATION SUMMARY

PATENT APPLICATION: US/09/800,770

DATE: 01/14/2003

TIME: 14:55:04

Input Set : A:\A68614-1.ST25.txt

Output Set: N:\CRF4\01142003\I800770.raw

L:2546 M:258 W: Mandatory Feature missing, <223> Blank for SEQ#:27,Line#:2543
L:4228 M:258 W: Mandatory Feature missing, <223> Blank for SEQ#:47,Line#:4225